Analysis of feature relevance using an image quality index applied to digital mammography

Costa, Arthur C.; Barufaldi, Bruno; Borges, Lucas R.; Biehl, Michael; Maidment, Andrew D. A.; Vieira, Marcelo A. C.

Published in:
SPIE Medical Imaging 2019

DOI:
10.1117/12.2512975

Publication date:
2019

Analysis of feature relevance using an image quality index applied to digital mammography

Arthur C. Costa[a,c], Bruno Barufaldi[b], Lucas R. Borges[a], Michael Biehl[c], Andrew D. A. Maidment[b], and Marcelo A. C. Vieira[a]

[a]Dept. of Electrical and Computer Engineering, University of São Paulo, São Carlos, Brazil
[b]Dept. of Radiology, University of Pennsylvania, Philadelphia, USA
[c]Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence, University of Groningen, Groningen, The Netherlands

ABSTRACT

In previous work, we investigated the application of the normalized anisotropic quality index (NAQI) as an image quality metric for digital mammography. The initial assessment showed that NAQI depends not only on radiation dose, but also varies based on image features such as breast anatomy. In this work, these dependencies are analyzed by assessing the contribution of a range of features on NAQI values. The generalized matrix learning vector quantization (GMLVQ) was used to evaluate feature relevance and to rank the imaging parameters and breast features that affect NAQI. The GMLVQ uses prototype vectors to segregate and to analyze the NAQI in three classes: (1) low, (2) medium, and (3) high NAQI values. We used Spearman's correlation coefficient (ρ) to compare the results obtained by the GMLVQ method. The GMLVQ was trained using 6,076 clinical mammograms. The statistical analysis showed that NAQI is dependent on several imaging parameters and breast features; in particular, breast area (ρ = -0.65), breast density (ρ = 0.62) and tube current-exposure time product (mAs) (ρ = 0.56). The GMLVQ results show that the most relevant parameters that affect the NAQI values were breast area (approx. 31%), mAs (approx. 24%) and breast density (approx. 15%). The GMLVQ method allowed us to better understand the NAQI results and provide support for the use of this metric for image quality assessment in digital mammography.

Keywords: digital mammography, generalized matrix learning vector quantization, normalized anisotropic quality index.

1. INTRODUCTION

The Anisotropic Quality Index (AQI) was introduced by Gabarda and Cristóbal to evaluate image quality of natural images regardless of any prior knowledge or reference (i.e., ground-truth).[1] This no-reference index was initially proposed to evaluate the performance of distortion-specific noise types, such as JPEG, white noise, and Gaussian blur.[1] We have been evaluating the use of this index to assess the quality of digital mammograms.

We have demonstrated and validated the use of the normalized AQI (NAQI) for assessing the quality of clinical mammograms.[2–4] NAQI can be used for image quality assessment in clinical practice where subject ground-truth does not exist.[2] The NAQI is comparable to well-established full-reference metrics when assessing the quality of digital breast phantoms and clinical mammograms in terms of noise and spatial resolution.[2]

We have also investigated the influence of breast anatomy on the NAQI values.[3] This particular study used a digital anthropomorphic phantom to simulate a range of breast anatomies. Tests were performed to analyze the effect of breast composition and radiation dose on the quality index.[3] Results showed that the metric is more sensitive to changes in radiation dose than to differences in the simulated breast anatomy.

Finally we evaluated NAQI performance using clinical mammograms.[4] The effect of radiation dose, breast density and compressed breast thickness on NAQI values was analyzed. The NAQI values demonstrated a negative correlation with compressed breast thickness and a positive correlation with average glandular dose and breast density.[4] However, a detailed analysis of the NAQI values is required to quantify the contribution of each imaging parameter and breast feature on image quality.

The current study presents a more detailed investigation of the effect of NAQI on various imaging parameters and breast features. We used the generalized matrix learning vector quantization (GMLVQ) to quantify the effect of each imaging parameter and breast feature on NAQI values. The GLMVQ method was used to calculate feature relevance (i.e., rank of features).

2. MATERIALS & METHODS

2.1 NAQI Analysis

A detailed description of NAQI analysis can be found in previous papers.[2–5] NAQI is calculated locally in small regions of interest (ROIs). In this work, the local NAQI values were calculated using a window size of 16 pixels, and the entropy direction was set to four angles ($\theta = [0°, 45°, 90°, 135°]$).[1,6] Due to the high contrast between the periphery of the breast and the image background, the local NAQI values located in the skin region were removed using the binary mask of the breast.[4] Finally, all local NAQI values in the inner region of the breast were averaged to provide a general index.

The NAQI analysis was performed using 6,076 clinical cranio-caudal mammograms. The mammograms were acquired using four Hologic Selenia Dimensions mammography systems (Hologic Inc., Bedford, MA) installed at the Hospital of the University of Pennsylvania.[4] To eliminate the effect of image processing on the clinical mammograms, we used raw images (DICOM For Processing). The use of the clinical mammograms was approved by the institutional review board, and a HIPAA-compliant retrospective interpretation of prospectively acquired data was conducted.

2.2 Feature Selection

For this study, a total of five imaging parameters and three breast features were used for analysis (Table 1): patient age, tube voltage (kV), tube current-exposure time product (mAs), average glandular dose (AGD), incident air-kerma (IAK), compressed breast thickness (CBT), breast area (BA) and percentage of breast density (PD). PD and BA measurements were estimated using the open-source LIBRA software.[5] The remaining parameters were extracted from the DICOM header.

Table 1: Descriptive statistics of features from all clinical images

	Minimum	**Maximum**	**Mean**	**Sd. Deviation**
Age	25	88	55.98	10.81
kV	25	37	29.75	1.68
mAs	56	445	129.61	40.57
AGD (mGy)	0.54	5.07	1.36	0.45
IAK (mGy)	1.09	23.16	5.27	2.18
CBT (mm)	16	105	57.15	12.34
BA (cm^2)	31.27	540.40	187.64	95.56
PD (%)	0.001	90.32	14.82	10.58
NAQI	0.05	0.21	0.11	0.03

2.3 Correlation coefficient

Correlation is a statistical method used for evaluation of linear associations between two variables. The two main types of correlation coefficients are Pearson's product moment (parametric) and Spearman's rank, or Spearman's coefficient (ρ) (non-parametric). Correlation values range from -1 to 1. The closer the coefficient is to -1, the more the association is inversely proportional, and closer it is to 1, the more the dependency of the two variables is directly proportional. The distribution of the variables define which correlation type is valid. Pearson's correlation requires that the variables follow a normal distribution. Spearman's method is used for non-normal distributed variables and it is robust to outliers.[7]

Although the Shapiro-Wilk normality test is considered the best of its kind for distribution analysis, it fails to treat data with a large number of samples (i.e. >5000).[8] Therefore, the Kolmogorov-Smirnov (KS) test was used for this study. Besides the normality tests, some authors suggest assessing distribution measures, like skewness and kurtosis, for a more reliable indicator for large samples.[9]

2.4 Generalized Matrix Learning Vector Quantization

Among machine learning algorithms, there is a popular algorithm family known as learning vector quantization (LVQ), which was introduced by Kohonen as a simple, interpretable, and powerful classifier.[10] LVQ is a supervised learning algorithm that uses labeled data for training. The algorithm assigns the inputs to discrete categories represented by prototype vectors.[10] The prototypes mark the coordinates in the input feature space where the class region is localized and provide information on neighboring data. During the learning procedure, the input samples are presented iteratively; in each step, a parameter update occurs by moving the prototypes away or toward the sample based on its label. For classification, a new pattern is assigned to the class of the nearest prototype according to a selected distance measure.[11]

The generalized matrix LVQ (GMLVQ) is an extension of the LVQ, developed as a general metric learning approach that deals with feature relevance in the learning task. The relevance profile of the model is related to the amount of the contribution of each dimension in the classification. Considering that relevance of features can vary within the input data space, the GMLVQ algorithm modifies the distance measure to improve the classification performance and interpretability of results.[12] A full matrix of adaptive relevance is responsible for parameterizing the general Euclidean distance measure.[12]

Consider $\boldsymbol{\xi} \in \mathbb{R}^N$ is an unknown pattern, where N denotes the data dimensionality, and $\boldsymbol{w} \in \mathbb{R}^N$ is a prototype. The generalized distance for the GMLVQ is

$$d_{\boldsymbol{\Lambda}}\left(\boldsymbol{\xi}, \boldsymbol{w}\right) = \left(\boldsymbol{\xi} - \boldsymbol{w}\right)^T \boldsymbol{\Lambda} \left(\boldsymbol{\xi} - \boldsymbol{w}\right), \tag{1}$$

where $\boldsymbol{\Lambda}$ is an $N \times N$ matrix of adaptive relevance of each feature. For Equation (1) to be considered a meaningful distance measure, $\boldsymbol{\Lambda}$ should be non-negative and symmetric, following certain specifications.[12] These conditions guarantee that the measure corresponds to a squared Euclidean distance in the transformed space. To achieve these requirements, $\boldsymbol{\Lambda}$ is decomposed into $\boldsymbol{\Lambda} = \boldsymbol{\Omega}^T\boldsymbol{\Omega}$, where $\boldsymbol{\Omega} \in \mathbb{R}^{MxN}$, with $M = N$ in the simplest case. This decomposition yields $\boldsymbol{u}^T\boldsymbol{\Lambda}\boldsymbol{u} = \boldsymbol{u}^T\boldsymbol{\Omega}^T\boldsymbol{\Omega}\boldsymbol{u} = (\boldsymbol{\Omega}^T\boldsymbol{u})^2 \geq 0$ for all $\boldsymbol{u}$.[12] Thus, $d_{\boldsymbol{\Lambda}}$ in Equation (1) can be rewritten as

$$d_{\boldsymbol{\Lambda}}\left(\boldsymbol{\xi}, \boldsymbol{w}\right) = \left[\left(\boldsymbol{\xi} - \boldsymbol{w}\right)^T \boldsymbol{\Omega}^T\right]\left[\boldsymbol{\Omega}\left(\boldsymbol{\xi} - \boldsymbol{w}\right)\right] = \left[\boldsymbol{\Omega}\left(\boldsymbol{\xi} - \boldsymbol{w}\right)\right]^2. \tag{2}$$

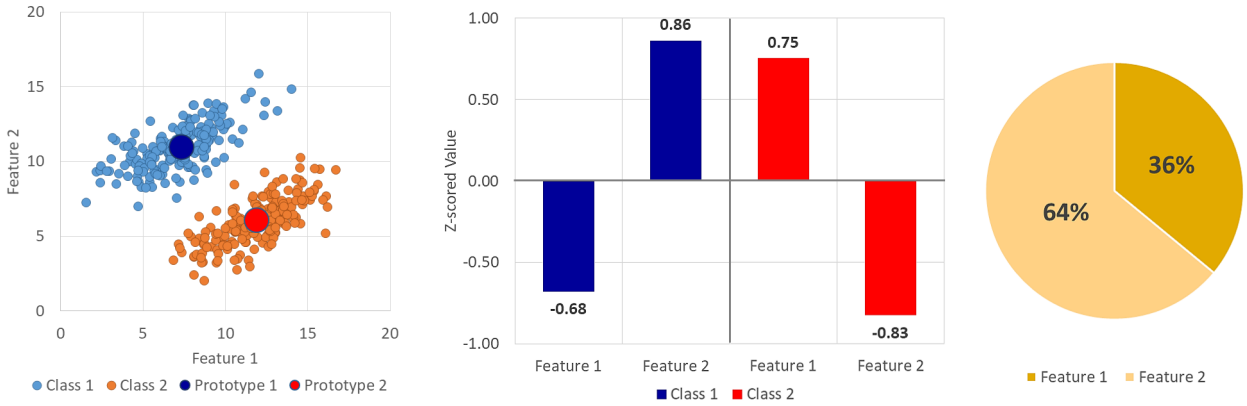
In the simplest version of GMLVQ, $\boldsymbol{\Omega}$ is a global matrix responsible for the data space transformation.[12] The algorithm iteratively updates the matrix parameters and prototype vectors by means of stochastic gradient descent with respect to the following cost function:[13]

$$E_{GMLVQ} = \sum_{i=1}^{P} \Phi(\mu_i^{\Lambda}) = \sum_{i=1}^{P} \Phi\left(\frac{d_{\boldsymbol{\Lambda}}(\boldsymbol{w}_J, \xi_i) - d_{\boldsymbol{\Lambda}}(\boldsymbol{w}_K, \xi_i)}{d_{\boldsymbol{\Lambda}}(\boldsymbol{w}_J, \xi_i) + d_{\boldsymbol{\Lambda}}(\boldsymbol{w}_K, \xi_i)}\right), \tag{3}$$

where P is the number of samples, Φ is the identity function, and $d_{\boldsymbol{\Lambda}}(\boldsymbol{w}_J, \boldsymbol{\xi})$ and $d_{\boldsymbol{\Lambda}}(\boldsymbol{w}_K, \boldsymbol{\xi})$ are the distances of the pattern $\boldsymbol{\xi}$ to the closest correct and incorrect prototype, respectively.[14] By the end of the training procedure, $\boldsymbol{\Lambda}$ diagonal values should represent the relevance of the N original features. Therefore, each element of the matrix is divided by the sum of the diagonal after each iteration, ensuring that $\sum_i \Lambda_{ii} = 1$.[12] For more information on the mathematical background, see previous publications of the method[12] and related algorithms.[15] One significant difference from similar algorithms, like Principal Component Analysis (PCA),[15] is the fact that GMLVQ takes the class labels into account, which makes it more robust at separating intricate data from different classes.[11]

To illustrate how the relevance matrix is interpreted, consider the following example (Figure 1). Assume that two classes of points are generated in a 2D space with normal distributions. Each point has a value on the

horizontal axis corresponding to feature 1, and another value on the vertical axis corresponding to feature 2. The clusters can be rotated in the input data space to generate many different cases. However, we shall consider one orientation of the groups (Figure 1a). The GMLVQ prototypes for the case are plotted over the input data space in Figure 1a. Their z-score values are shown Figure 1b. The relevance of each feature expressed by $\mathbf{\Lambda}$ diagonal is represented in Figure 1c.



(a) (b) (c)

Figure 1: (a) Groups of points belonging to two different classes with the corresponding prototypes obtained by the GMLVQ; (b) Z-score of the prototypes; (c) Feature relevance.

Depending on the orientation of the data sets to the feature axes, the relevance will lean more towards one feature or the other, meaning classes are more distinguishable by one feature. The more the two sets are horizontally oriented, the more relevant feature 2 is to the classification task, meaning that the problem gets closer to a scenario where the clusters can be separated exclusively by this feature. The opposite happens when the data sets are vertically oriented.

The original GMLVQ algorithm is not suitable for regression tasks. Therefore, we organized the data to fit a classification problem by dividing the entire mammogram set into three subsets based on average NAQI values. We equally divided the entire range of NAQI values ($[0.05, 0.21]$) into three equal classes; class 1 represents mammograms with the lowest NAQI values in the range of $[0.05, 0.10)$; class 2 represents the images with intermediate values of NAQI in range of $[0.10, 0.12)$; and class 3 represents the mammograms with the highest NAQI values in the range of $[0.12, 0.21]$.

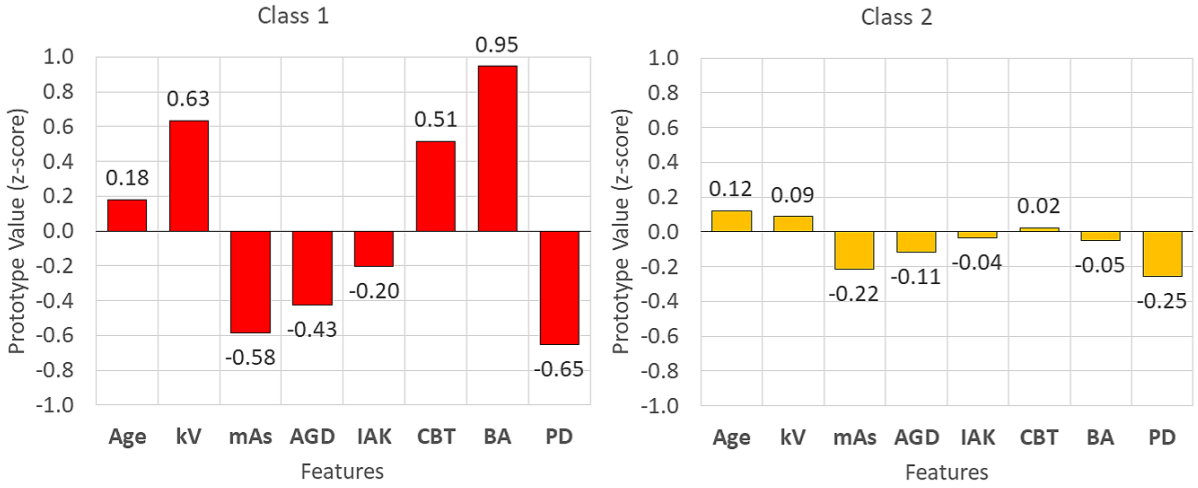
3. RESULTS

Analysis showed non-normal distributions of the imaging parameters and breast features, as well as NAQI (i.e., all variables failed the KS normality test). Thus, we analyzed the distributions using non-parametric methods. Spearman's rank correlation coefficient (ρ) was chosen to assess the linear relationship between variables (Table 2). The results were statistically significant at the 0.01 level (2-tailed test).

Table 2: Spearman's ρ correlation coefficient calculated pairwise between NAQI and features

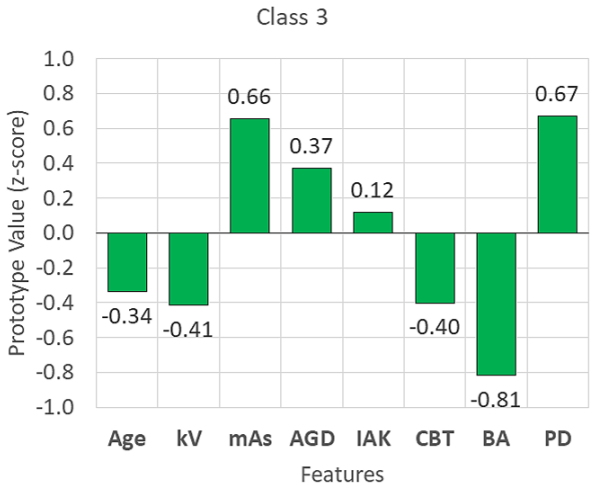
	Age	**kV**	**mAs**	**AGD**	**IAK**	**CBT**	**BA**	**PD**
ρ	-0.16	-0.33	0.56	0.40	0.18	-0.35	-0.65	0.62

Table 2 shows that mammograms with high NAQI are acquired with high dose (high mAs, AGD and IAK values) and dense breasts. Large and fatty breasts (i.e., low PD, and high CBT and BA values) tend to generate mammograms with lower NAQI values.



(a)

(b)



(c)

Figure 2: Prototypes values normalized in z-score for all classes: 1(a), 2(b) and 3(c)

The GMLVQ algorithm converged for the three prototypes shown in Figure 2. Each feature value is z-score normalized. In addition, the relevance of each feature is presented in Figure 3.

The prototypes generated by class 1 and class 3 show a clear inversion of the analyzed features (Figure 2). Thus, there is a clear distinction between the vectors associated with the imaging parameters and breast features, and NAQI. As indicated by Figure 2a, low NAQI is associated with thicker and fattier breasts. Figure 2c indicates that high NAQI is associated with thinner and denser breasts.

Figure 3 shows that BA (30.9%) is the most relevant feature for the NAQI analysis. As expected, NAQI is also sensitive to radiation dose. The mAs was classified as the second most relevant with 24.1%. PD also affects the NAQI values, which represents the third most relevant feature for the NAQI analysis (14.7%). Note that the impact of these features in the NAQI classification analysis also matches with the results of correlation coefficients, considering that they had the strongest correlation with NAQI (Table 2).

Figure 4 illustrates the colormaps of NAQI values before (middle) and after (right) skin segmentation from two different patients. As we mentioned previously, for the NAQI classification we used the average NAQI values after the skin segmentation. Figure 4a shows one example of a breast 65 mm thick, with BA = 172.05 cm^2, and

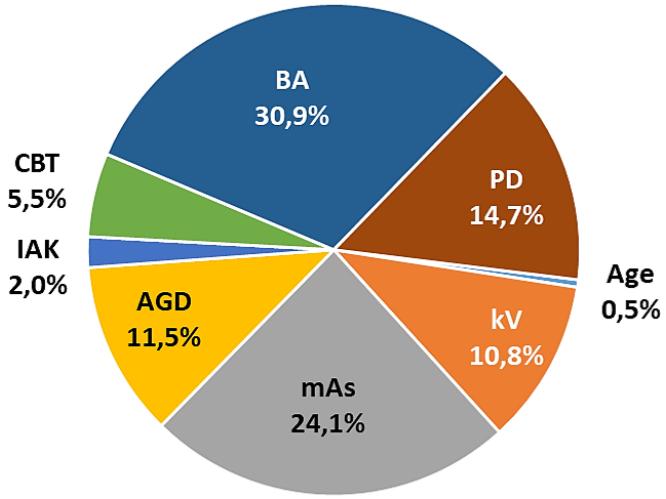


Figure 3: Feature relevance corresponding to the diagonal elements of matrix $\mathbf{\Lambda}$

PD = 8.92%. Figure 4b shows one example of a breast 69 mm thick, with BA = 93.97 cm^2, and PD = 19.88%. Comparing both mammograms, the average NAQI increased 70% in Figure 4b because of the differences in the imaging parameters and breast features. The GMLVQ method classified Figure 4a as class 1 and Figure 4b as class 3.

Similarly, Figure 5 illustrates the colormaps of NAQI values before (middle) and after (right) skin segmentation from two patients. In this example, both mammograms were *matched* in terms of CBT, BA, PD, and AGD. The differences in the breast parameters were less than 1%. Comparing both mammograms, the difference in average NAQI was only 6%. Both mammograms were classified as class 3.

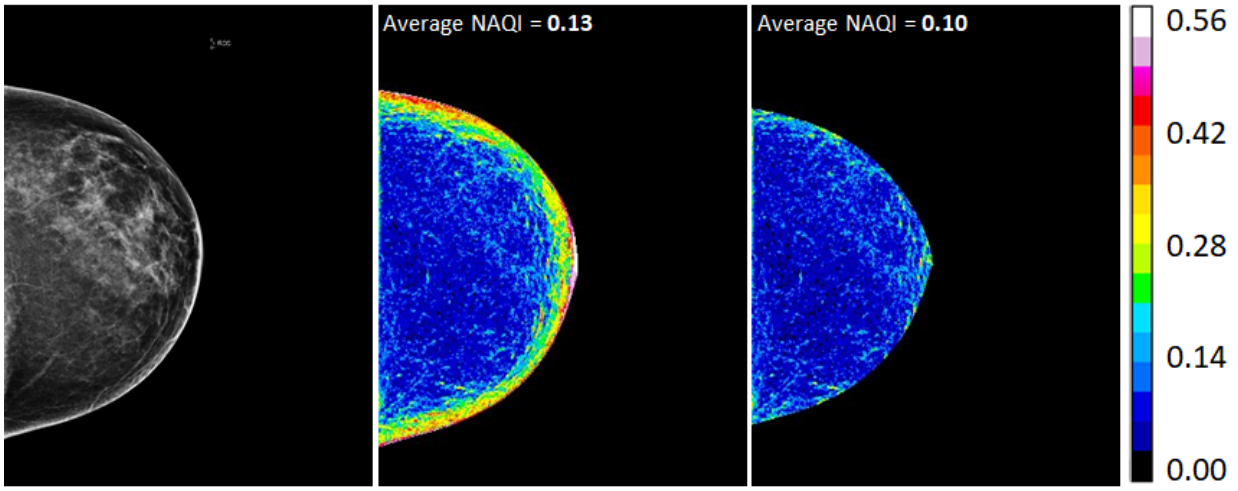
4. DISCUSSIONS

The NAQI prototypes and relevance can be used to indicate the parameters that affect NAQI image quality assessment in DM. As illustrated in Figure 2, the prototypes of class 1 and class 3 are complementary, where all features are inversely proportional. Class 2 represents mammograms that are neither considered with a low nor high NAQI value. Figure 2b shows the prototype of class 2. As can be seen, the bar plot is "flat", representing parameter values close to their corresponding mean values. Further analysis is required to assess the images classified as class 2 objectively.

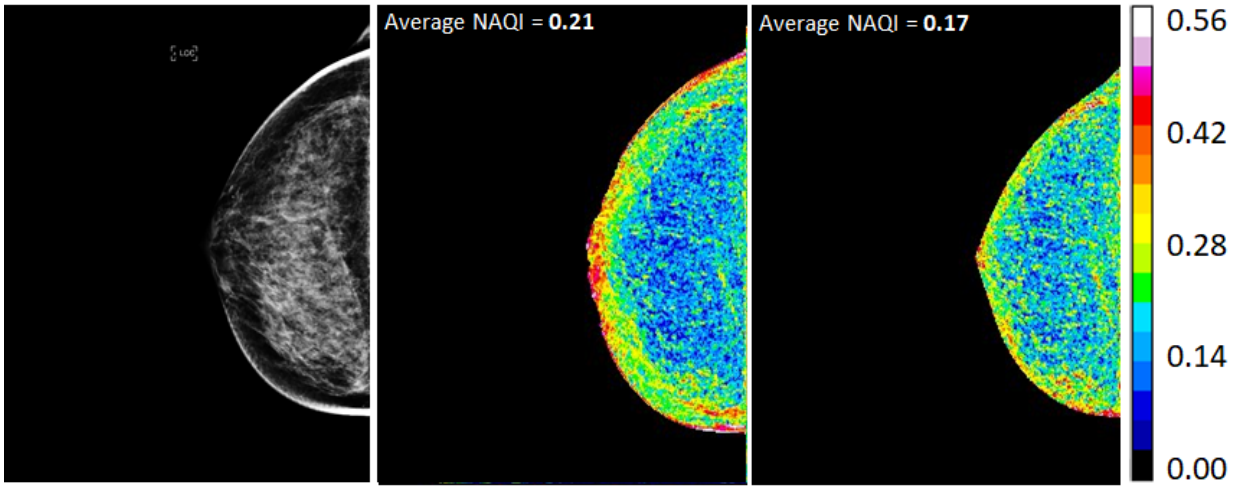
Using the GMLVQ method we were able to quantify the contribution (i.e., relevance) of each imaging parameter and breast feature to NAQI. As seen in Figure 3, NAQI is most sensitive to BA, mAs, PD. These results were consistent with previous work.[4] It is important to note that radiation dose levels are directly associated with mAs and image quality. Therefore, mAs relevance and correlation results with NAQI were as expected.

Figure 4a shows a scattered fibroglandular breast (PD = 8.92%) and Figure 4b shows a heterogeneously dense breast (PD = 19.88%). In terms of BA, Figure 4a represents an average breast (8% larger than the average BA - Table 1) and 4b represents a small breast (50% smaller than the average BA - Table 1). The GLMVQ algorithm classified these mammograms in classes 1 and 3 (approx. 70% difference in average NAQI), based on the imaging parameters and breast features. To confirm the GLMVQ results, the imaging parameters and breast features were matched in Figure 5. The GLMVQ algorithm classified both mammograms in class 3, showing approximately 6% difference in average NAQI.

Figure 6a shows an example of an extremely dense breast (PD = 44.51%). The x-ray attenuation is higher in the region where the density clusters are located. Thus, we expect lower image quality in these regions. The local NAQI colormap (middle) shows low values of local NAQI located in the region of dense clusters and high values elsewhere. This example also illustrates that in denser breasts, the non-dense regions are responsible for raising the NAQI average. Even though the percentage of dense tissue tends to be less than 20% on average,[16]



(a)



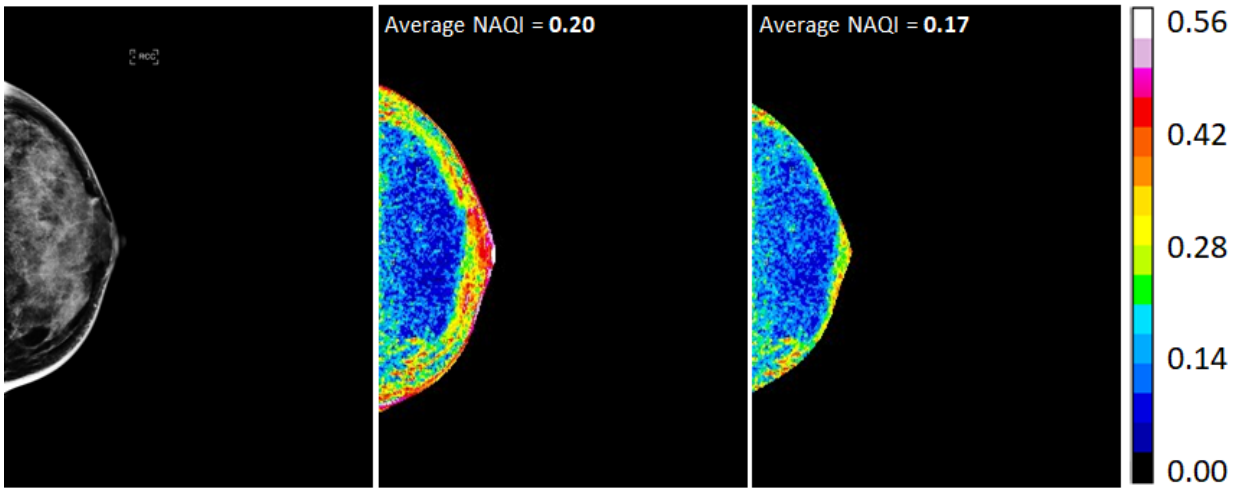
(b)

Figure 4: Original mammogram (left), colormap of the local values of NAQI (middle), and colormap of the local values of NAQI after skin segmentation (right). These images represent examples of the NAQI classification in (a) class 3 and (b) class 1.

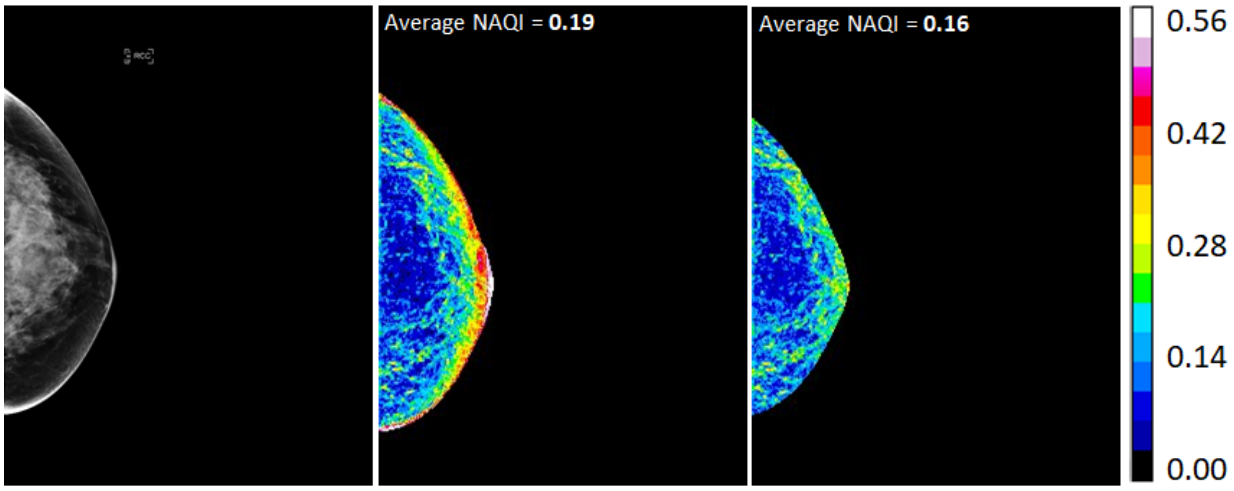
when the NAQI is expressed as a mean, it does not capture the particularities of each region. For future work, we will segment the NAQI values in dense cluster regions ($NAQI_d$).

Figure 6b shows an example of a fatty and thick breast (PD = 11.86% and CBT = 77 mm). The x-ray attenuation is higher for thicker breasts. Thus, we expect lower image quality for these cases. The local NAQI colormaps show lower local NAQI values within the inner region.

As we mentioned in previous work,[4] the local NAQI values that are found close to the breast skin ($NAQI_s$) are extremely high because of the high contrast between the periphery of the breast and the image background. We used breast masks[4] to remove the $NAQI_s$ values. These binary masks were eroded based on the chest wall-to-nipple distance (d). Then, we multiplied the eroded mask by the values of the local NAQI. The results of the skin segmentation are shown in Figure 6. Note that the segmentation of $NAQI_s$ is not accurate because the amount of erosion was chosen arbitrarily (10% of d). The white arrows highlight some of the local NAQI values close to the skin that should have been removed (Figure 6). Thus, an adaptive segmentation of the $NAQI_s$ is required. For future work, we will use the maximum entropy to threshold and segment $NAQI_s$ values effectively.



(a)



(b)

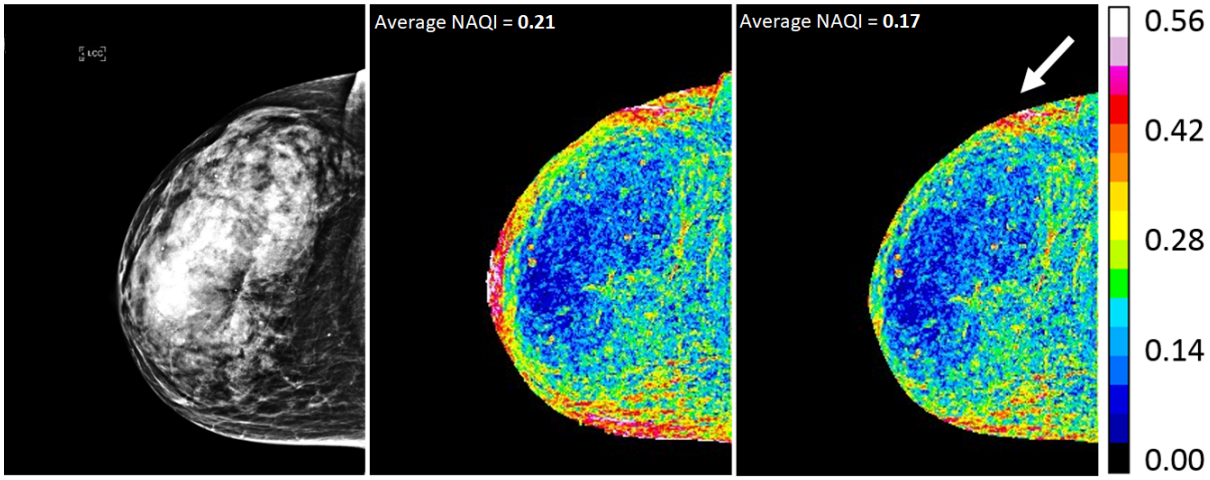
Figure 5: Two examples of mammograms that were matched in terms of CBT, BA, PD, and AGD, both classified as class 3. Original mammogram (left), colormap of the local values of NAQI (middle), and colormap of the local values of NAQI after skin segmentation (right).

Although Figures 6a and 6b show processed mammograms (left), the NAQI values were calculated using raw images. In previous publication,[4] we showed that the image processing of clinical mammograms affects the NAQI values. Note that Figure 6b shows a breast-whithin-breast artifact.[17] For future work, we will also perform the NAQI analysis using processed images to evaluate breast-whithin-breast artifacts.

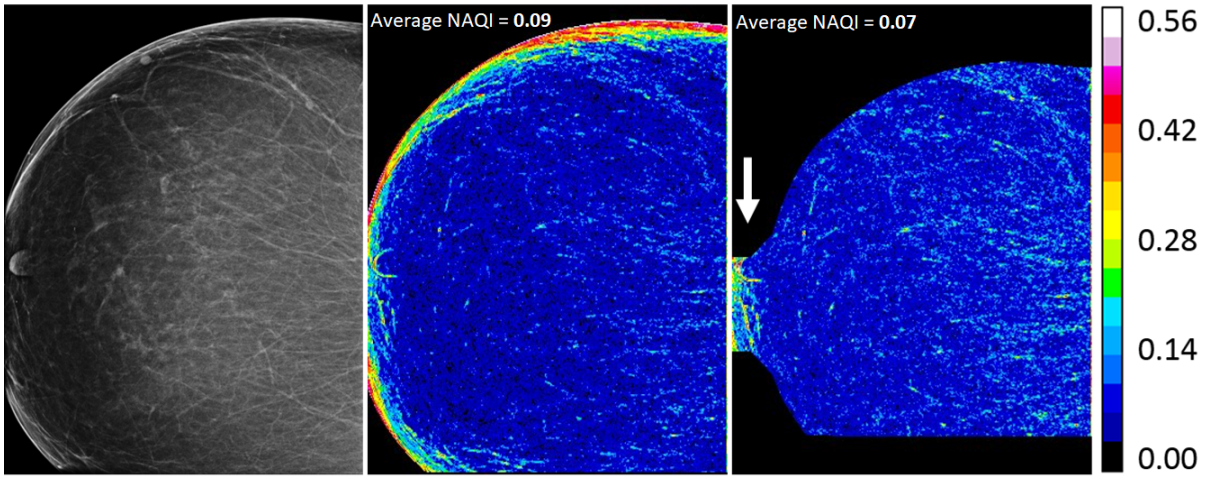
5. CONCLUSION

This study presents a detailed investigation of the effect of NAQI on various imaging and breast parameters. The GMLVQ method supports the NAQI obtained in previous studies, as well as enhances the use of NAQI as a metric for image quality assessment in DM. The average NAQI is strongly dependent on several imaging and breast parameters, such as BA, mAs and PD.

The prototypes of class 1 (low NAQI values) and class 3 (high NAQI values) presented an opposite relationship with the five image parameters and the three breast features. Low NAQI is associated with larger fattier breasts



(a)



(b)

Figure 6: Examples of failures in skin segmentation. Original mammogram (left), colormap of the local values of NAQI (middle), and colormap of the local values of NAQI after skin segmentation (right). These images represent examples of the NAQI classification in (a) class 3 and (b) class 1.

(e.g. Figure 2a), and high NAQI is associated with smaller denser breasts (e.g. Figure 2b). Because NAQI is expressed as an average of local NAQI in the inner region of the breast, the contribution of each imaging parameter and breast feature in the average NAQI can be misrepresented. NAQI can be considered as a more valuable image quality metric when evaluated locally, using segmented regions of the breast (e.g., local average NAQI at the region located at dense clusters).

For future work, we will conduct a more in-depth investigation of our data using mediolateral oblique exams. We will also investigate how much the density clusters affect the NAQI classification. We believe that a superior NAQI segmentation will improve the responses of the GMLVQ algorithm in classes 1-3. For that purpose, we will calculate an adaptive threshold. Finally, the GLMVQ algorithm will be used to train and test the segmented NAQI values (i.e., NAQI_s, NAQI_d, and NAQI_f) in different breast regions.

ACKNOWLEDGMENTS

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. It was also supported by Erasmus +, within the Key Action 1 Mobility of Individuals program, the São Paulo Research Foundation (FAPESP) grant 2018/19888-5, Susan G. Komen Foundation Grant IIR13264610, NIH/NCI Grant 1R01CA154444, NIH PROSPR Network Grant U54-CA163313, and the Burroughs-Wellcome Fundation (IRSA 1016451).

REFERENCES

[1] Gabarda, S. and Cristóbal, G., "Blind image quality assessment through anisotropy," *JOSA A* **24**(12), B42–B51 (2007).

[2] de Oliveira, H. C., Barufaldi, B., Borges, L. R., Gabarda, S., Bakic, P. R., Maidment, A. D., Schiabel, H., and Vieira, M. A., "Validation of no-reference image quality index for the assessment of digital mammographic images," in [*Medical Imaging 2016: Image Perception, Observer Performance, and Technology Assessment*], **9787**, 978713, International Society for Optics and Photonics (2016).

[3] Barufaldi, B., Borges, L. R., Vieira, M. A., Gabarda, S., Maidment, A. D., Bakic, P. R., Pokrajac, D. D., and Schiabel, H., "The effect of breast composition on a no-reference anisotropic quality index for digital mammography," in [*International Workshop on Digital Mammography*], 226–233, Springer (2016).

[4] Barufaldi, B., Borges, L. R., Bakic, P. R., Vieira, M. A., Schiabel, H., and Maidment, A. D., "Assessment of automatic exposure control performance in digital mammography using a no-reference anisotropic quality index," in [*Medical Imaging 2017: Image Perception, Observer Performance, and Technology Assessment*], **10136**, 101360U, International Society for Optics and Photonics (2017).

[5] Keller, B. M., Nathan, D. L., Wang, Y., Zheng, Y., Gee, J. C., Conant, E. F., and Kontos, D., "Estimation of breast percent density in raw and processed full field digital mammography images via adaptive fuzzy c-means clustering and support vector machine segmentation," *Medical physics* **39**(8), 4903–4917 (2012).

[6] Gabarda, S., Cristóbal, G., and Goel, N., "Anisotropic blind image quality assessment: Survey and analysis with current methods," *Journal of Visual Communication and Image Representation* **52**, 101–105 (2018).

[7] Mukaka, M. M., "A guide to appropriate use of correlation coefficient in medical research," *Malawi Medical Journal* **24**(3), 69–71 (2012).

[8] Razali, N. M., Wah, Y. B., et al., "Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests," *Journal of statistical modeling and analytics* **2**(1), 21–33 (2011).

[9] Kim, H.-Y., "Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis," *Restorative dentistry & endodontics* **38**(1), 52–54 (2013).

[10] Kohonen, T., "Learning Vector Quantization," in [*Self-organizing maps*], 203–217, Springer (1997).

[11] Schneider, P., Biehl, M., and Hammer, B., "Adaptive relevance matrices in Learning Vector Quantization," *Neural computation* **21**(12), 3532–3561 (2009).

[12] Schneider, P., Biehl, M., and Hammer, B., "Distance learning in discriminative vector quantization," *Neural computation* **21**(10), 2942–2969 (2009).

[13] Sato, A. and Yamada, K., "Generalized Learning Vector Quantization," in [*Advances in neural information processing systems*], 423–429 (1996).

[14] Schneider, P., [*Advanced methods for prototype-based classification*], Atto Producties Europe (2010).

[15] Smith, L. I., "A tutorial on principal components analysis," tech. rep. (2002).

[16] Yaffe, M., Boone, J., Packard, N., Alonzo-Proulx, O., Huang, S.-Y., Peressotti, C., Al-Mayah, A., and Brock, K., "The myth of the 50-50 breast," *Medical physics* **36**(12), 5437–5443 (2009).

[17] Pisano, E. D., Cole, E. B., Hemminger, B. M., Yaffe, M. J., Aylward, S. R., Maidment, A. D., Johnston, R. E., Williams, M. B., Niklason, L. T., Conant, E. F., et al., "Image processing algorithms for digital mammography: a pictorial essay," *Radiographics* **20**(5), 1479–1491 (2000).